ANGRY, HAPPY, SAD OR PLAIN NEUTRAL? THE IDENTIFICATION OF VOCAL AFFECT BY HEARING-AID USERS

Cécile Pereira
Speech, Hearing and Language Research Centre
Macquarie University

ABSTRACT - This paper reports a speech perception study on the identification of vocal affect by hearing aid users. The performance of a group of 40 normally-hearing subjects is compared to that of a group of 39 post-lingually deafened subjects using hearing-aids. All subjects are adult native speakers of English. Results indicate that there are major differences between the two groups. Scores for overall identification of affect are 85% for the normally hearing listeners and 65% for the hearing-aid users. Patterns of confusion in the identification of emotions are similar in the two groups, but there is a greater degree of confusion in the hearing-aid users. There is a significant negative correlation between hearing-loss and the correct identification scores.

INTRODUCTION

In social interactions, understanding the mood and attitude of a speaker is very important for understanding the meaning of a spoken message. Mood and attitude can be communicated verbally, but more often than not they are communicated non-verbally through body language, facial expressions and voice. This emotional meaning expressed non-verbally may reinforce the verbal message, add to it or contradict it. For the hearing-impaired, who have difficulties in understanding the verbal content of a spoken message, the perception of the emotion and attitude as expressed by the tone of voice is even more important socially than for normally hearing people, especially if visual cues are not available. For the hearing-impaired the issue is not just a question of "It ain't what you say, but the way you say it", but "It ain't what you say, but the way I hear it".

Few published studies exist to date which examine the perception of vocal affect by the hearing impaired. Anne-Marie Öster and Arne Risberg (1986) showed that hearing-impaired children and adults fitted with hearing-aids had difficulty in identifying the mood of a speaker in test sentences, particularly confusing happy, angry and sad. In his research, David House (1989, 1990 (a,b), 1991) replicated these findings. He also found that the confusions formed two main patterns: happiness was confused with anger, and sadness was confused with neutrality. House (1991) working on a sample of 29 listeners with moderate to severe hearing losses found that performance did not correspond to the degree of hearing loss, which seems surprising, but is nevertheless corroborated by the findings of Most et al (1993) who worked on a sample of 24 severe to profound listeners. However, Most et al suggest that with a group including subjects with better residual hearing, it would be more likely that one could observe correlations between loss and correct identification of emotion.

This study addresses the question as to whether adults with normal hearing and those who are post-lingually deafened and are using hearing-aids perceive equally well the expression of mood in the speech of others.

METHOD

Subjects

All subjects were adult native speakers of English.

The hearing-aid users: The experimental group consisted of 39 post-lingually deafened subjects with mild to severe degrees of sensorineural hearing loss in one or both ears. The mean age for this group was 68 years, with a range of 46 to 83 years and a standard deviation of 9 years. There were 15 women and 24 men. The mean 3 frequency average loss for their better ear was 39 dBHL, with a range from 13

dBHL to 95dHL (the average was based on the loss at 500Hz, 1000Hz and 2000Hz).

The control group: Subjects in the control group had normal hearing bilaterally. In many experiments where experimental groups are compared to a control group, the groups are matched for age. This is not the case in this study because to compare impaired hearing with normal hearing, it was necessary for the control group to be in an age range where there is no impairment due to age. Therefore control subjects ranged in age from 22 to 58 years, with a mean age of 42 years and a standard variation of 9 years. There were 16 men and 24 women.

Corpus

Two actors, one male aged 50, and the other, female aged 40, both native speakers of English, were asked to portray given emotions by reading specially prepared scripts. This was done for three different emotions: happiness, anger, sadness, and for a neutral state, serving as a reference for comparisons. The scripts for these four different affect states were based on situations reported by David House (1990) for Swedish subjects. Each situation was meant to provide a context for the production of a semantically neutral phrase, the number "2510" (uttered "two thousand five hundred and ten") and a semantically neutral sentence, "I'm going to move house". For the number "2510" the situations were as follows: Neutral - the number of books in the library. Happy - the dollar total of an unexpected tax rebate; Angry - the dollar total of an unexpected tax debt; and Sad, the loss of a wallet containing the unexpected tax rebate. For the sentence "I'm going to move house", the situations were as follows: Neutral - reporting a change of address to the Electoral Office. Happy - telling friends about the acquisition of a new house. Angry - telling friends about the landlord who has drastically raised the rent of a house in disrepair, and Sad, telling friends the news of a separation and its consequences.

Tape recordings were made of the two actors each acting the 8 different scripts. Several takes of each rendition were recorded, to allow for selection of the most representative takes later. For anger, the actors were asked to portray two forms of anger - "hot" or "red" anger, and "cold" or "repressed" anger. This was done because the two forms are acoustically quite different, but are both commonly used.

Stimuli tape

Based on the selection of 10 judges, the best two productions were retained for each combination of emotion, utterance, and actor. For anger, four productions were retained, two for hot anger, another two for cold anger. These selected productions were repeated 3 times, thus creating 120 presentations for the final experimental stimuli. These were organized in four groups of 30. The first 30, for the utterance "I'm going to move house" said by the male actor, the second 30, for the utterance "2510" said by the male actor also, and similarly for the two utterances spoken by the female actor. Within each group, the tokens were presented in random order.

Another recording was made, in which the average intensity of each token was normalised, thus removing the intensity cue as a means of identification.

Procedure

Each subject was tested in two experimental sessions, the first session for the non-normalised recording and the second session for the recording where the intensity had been normalised. Using a forced choice paradigm, subjects were asked to categorise each stimulus as one of four possible emotional states, neutral, angry, sad or happy, by ticking a box on a questionnaire. Subjects were asked to select anger, whether they heard cold or hot anger.

Subjects in the control group were tested with headphones. Subjects in the experimental group were tested in a sound treated room in a free field. At the end of their session, an aided audiogram was taken. The NAL/Arthur Boothroyd monosyllabic single word speech discrimination test was administered to the subjects who did not have such record on their file.

RESULTS

Control group results

Analysis of the results for the control group indicate 84% correct identification for Experiment 1 (no normalisation of intensity), with a 6% standard deviation, and 85% correct identification for Experiment 2 (intensity normalised) with a 7% standard deviation. The percentage of correct identification across tokens, repetitions, speakers, utterances and emotions hardly varies from one experiment to another indicating that for this group the normalisation of intensity did not make a difference.

There was a significant difference in the rate of identification according to the emotion, thus indicating that some emotions were significantly more difficult to identify. Sadness was the easiest to correctly identify (95% and 96% correct identification), followed by hot anger (88% and 84%), happiness (83% and 86%), neutrality (78% and 78%), and cold anger (74% and 79%). The main pattern across both experiments therefore seems to be, from most to least easily recognised: sadness, hot anger and happiness, neutrality and cold anger.

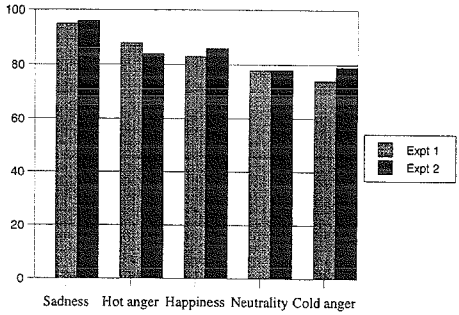

Figure 1. Percentages of correct identification per emotion for the control group.

Patterns of confusion are such that neutrality is often confused with sadness (21% in both experiments), cold anger with neutrality (15% and 14% in Experiment 1 and Experiment 2 respectively), happiness with neutrality (14% and 12%), and hot anger with happiness (10% and 14%).

	RESPONSE			
INTENDED	anger	happiness	neutrality	sadness
cold anger	74% 79%	10% 7%	15% 14%	2% <1%
hot anger	88% 84%	10% 14%	2% 2%	<1% 0%
happiness	2% 2%	83% 86%	14% 12%	<1% <1%
neutrality	<1% <1%	<1% <1%	78% 78%	21% 21%
sadness	<1% <1%	<1% 1%	3% 3%	95% 96%

Table 1. Confusion matrix for the control group and the two experiments (first line of percentages: Experiment 1; second line: Experiment 2).

There was no significant correlation between age and percentage of correct identification. There was also no significant correlation between sex and percentage of correct identification.

Hearing-aid user group results

Analysis of the results for this group indicate 65% correct identification for Experiment 1, with an 8% standard deviation, and 64% correct identification for Experiment 2, with a 11% standard deviation. This represents a difference of about 10% in comparison with the control group, and more individual variation.

	Expt 1	Expt 2
NH	84% ± 6%	85% ± 7%
HA	65% ± 8%	64% ±11%

Table 2. Correct identification scores for the control group and the hearing-aid user group.

The rate of correct identification per emotion is similar between both experiments, suggesting that the normalisation of intensity might not have had any influence. However, the results for the hearing-aid users are less similar between experiments than they were for the control group.

As was the case for the control group, there was a significant difference in the rate of identification according to the emotion, thus indicating that some emotions were more difficult to identify. Unlike for the control group, however, there does not appear to be a strong pattern of most to least easily recognized emotion across both experiments, nor is there a pattern similar to that of the control group, with the exception of cold anger, which showed as the most difficult to identify emotion. For Experiment 1, hot anger was the easiest to correctly identify (82%), followed by sadness (79%), neutrality (69%), happiness (59%) and cold anger (38%). For Experiment 2, it is sadness (72%), followed by neutrality (67%), happiness (62%), hot anger (69%) and cold anger (48%).

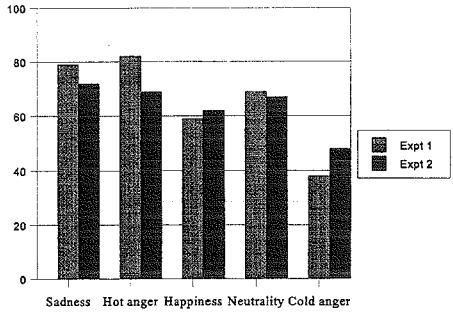

Figure 2. Percentages of correct identification per emotion for the hearing-aid users.

Patterns of confusion are such that neutrality is often confused with sadness (24% and 23% in Experiment 1 and Experiment 2 respectively), cold anger with neutrality (30% and 24%), happiness with neutrality (21% in both experiments), and hot anger with happiness (12% and 20%). These confusions are similar to those of the control group with regard to which emotion is mostly confused with which other, but different in terms of their amount (greater with the hearing-aid users). In addition, the hearing-aid users also confused cold anger with happiness (16% and 14%) and sadness (15% and 13%), happiness with hot anger (17% and 14%) and sadness with happiness (9% and 12%). The control group hardly seems to make these errors except for cold anger being confused with happiness (10% and 14%).

	RESPONSE			
INTENDED	anger	happiness	neutrality	sadness
cold anger	38% 48%	16% 14%	30% 25%	16% 14%
hot anger	82% 69%	12% 20%	5% 9%	1% 1%
happiness	17% 14%	59% 62%	21% 21%	3% 3%
neutrality	3% 6%	4% 4%	69% 67%	24% 23%
sadness	2% 6%	9% 12%	10% 10%	79% 72%

Table 3. Confusion matrix for the hearing-aid users and the two experiments (first line of percentages: Experiment 1; second line: Experiment 2).

There was no significant correlation between age and percentage of correct identification nor was there between sex and percentage of correct identification overall. However, in this group, men had higher correct identification scores overall. They did better particularly on hot anger (especially in the male speaker), sadness (especially in the female speaker), and neutrality.

Effect of hearing loss. Based on their audiogram (unaided) and their aided audiograms, three frequency averages were calculated for each subject, as well as an aided average at higher and lower frequencies. These averages were then correlated with the percentages of correct identifications. This analysis reveals a significant relationship between the subjects' hearing loss and their performance in the experiments. At 500Hz, 1000Hz and 2000Hz, unaided, the correlation was high: $r = -.62$, $p = < .0005$ (Experiment 1); $r = -.60$, $p = < .0005$ (Experiment 2). The correlation with the aided 3FA while less high is still important: $r = -.49$, $p = .001$ (Experiment 1); $r = .44$, $p = .005$ (Experiment 2). It is interesting to look at correlations with both aided and unaided audiograms and performance in the experiment because while the subjects did the experiment aided, thus justifying looking at their aided thresholds, it must be remembered that unaided thresholds give a more accurate representation of hearing handicap. The worse the hearing loss, the worse the audibility problem, not only in terms of thresholds, but in psycho-acoustic terms. The hearing-aids can only give subjects amplification.

	3FA (unaided)	3FA (aided)
Expt 1	$r = -.62^{***}$	$r = -.49^{***}$
Expt 2	$r = -.60^{***}$	$r = -.44^{**}$

Table 4. Correlations between hearing-loss and correct identification scores (*** = $p \leq .001$ ** = $p \leq .005$)

When correlating the subjects' speech discrimination scores on the NAL/Arthur Boothroyd list and their performance in the experiments, one also observed a significant relationship: $r = .38$ and $p = .017$; $r = .44$ and $p = .005$). As the more severe the hearing loss, the greater the difficulty in understanding speech usually is, this confirms the negative correlation found between hearing loss and the correct identification scores.

SUMMARY

This study confirms the results of the studies cited in the introduction which had showed that hearing-impaired people had difficulties correctly identifying emotion in speech. It was found that their correct identification rate is lower than that normally-hearing people, but that the errors they make are of a similar nature, neutrality confused with sadness, cold anger with neutrality, and hot anger with happiness. The confusion of hot anger with happiness is of course very important, especially as both

emotions carry very different meanings. Misinterpretation of hot anger for happiness and poor identification of cold anger can have negative consequences in social interaction. This study, unlike the ones cited, clearly shows a relationship between the amount of loss and speech discrimination ability and the ability to identify emotion. It also reminds us of the fact that there is more to hearing loss than loss of thresholds. Further research is currently undertaken examining the identification of emotion in cochlear implantees, and analysing the acoustic characteristics of anger, sadness and happiness. It is hoped that those results, together with the present findings, will bring us one step closer to understanding the differences in the perception of vocal affect between the hearing-impaired and normally hearing people.

REFERENCES

House, David (1989). Cues for the perception of mood in speech: implications for the hearing impaired. In *British Journal of Audiology,* 23, 171.

House, David (1990a). On the perception of mood in speech: implications for the hearing impaired. In *Working Papers* 36, 99-108, Lund University, Department of Linguistics.

House, David (1990b). On the perception of mood in speech. In *Phoneme 1*, Papers from Fonetic-90, the Fourth Swedish Phonetics Conference, Umeå, May 1990.

House, David (1991). On hearing impairments, cochlear implants and the perception of mood in speech. In *Proceeding of the Fifth National Phonetic Conference, Stockholm, May 1991.*

Most, Tova, Amatzia Weisel and Ariela Zaychik (1993). Auditory, visual and auditory-visual identification of emotions by hearing and hearing-impaired adolescents. *British Journal of Audiology* 27: 247-253.

Öster, Anne-Marie, and Arne Risberg (1986). The identification of the mood of a speaker by hearing impaired listeners. In the Speech Transmission Laboratory, *Quarterly Progress and Status Report 4*: 79–90, Royal Institute of Technology, Stockholm.